APPENDIX

F

Integrating Data and Filling Gaps: The Case of Household Travel

Budget constraints make it difficult for statistical agencies to garner sufficient resources to launch new data collection programs that are responsive to changing policy concerns and at the same time maintain and improve needed data series from the past. One way to free up resources for new or modified data collection is to integrate two or more existing data systems into a more cost-effective combined system. Even when data integration does not result in net cost savings, it can still be useful to undertake if the combined data are relevant for a wider range of analyses. Sometimes full integration is not possible or sensible, but partial links among data systems, achieved through such means as the use of consistent definitions for key variables, can significantly enhance their analytical power. Finally, efforts to relate multiple data systems will often identify important gaps that none of them currently fills.

To develop examples of possibilities for linking and integrating transportation data sources that BTS might usefully explore with other relevant agencies, we reviewed surveys that provide data on household transportation. Information on household travel, taking account of all transportation modes, is critical for many important transportation policy concerns, including access, safety, direct costs to the household sector, and indirect costs in terms of energy use, environmental effects, and economic productivity.

The two most important national surveys of household transportation are the American Travel Survey (ATS) sponsored by BTS and the Nationwide Personal Transportation Survey (NPTS) sponsored by the Office of Highway Information Management (OHIM) in the Federal Highway Administration (FHWA). The decennial census long-form sample, the Consumer Expenditure Survey (CEX) of the Bureau of Labor Statistics (BLS), and the Residential Transportation Energy

Consumption Survey (RTECS) of the Energy Information Administration (EIA) also provide relevant data. (EIA recently discontinued RTECS because of budget reductions.)

We drew two main conclusions from this review. First, there appears to be an opportunity to develop a more cost-effective data collection system for household travel by integrating the ATS and the NPTS. Second, looking across all of the existing surveys, there appear to be important data gaps that should be filled.

INTEGRATING THE ATS AND THE NPTS

The NPTS, which is currently conducted on the same 5-year cycle as the ATS, is designed to provide data on daily household travel patterns. The sample includes about 22,000 households, who are asked about trips during a specified travel day. They are also asked about longer trips (75 or more miles) over the previous two weeks, but these data are not adequate for purposes of analysis given the short reference period and small sample size. The NPTS sample size also limits the geographic areas for which estimates can be published to the United States, urban areas as a whole, rural areas as a whole, and groups of cities categorized by population size. A few states and metropolitan planning organizations (MPOs) pay for additional samples for their areas. (Many states, MPOs, and localities also conduct their own travel surveys independently.)

The ATS, under its current design, includes a sample of 80,000 households, who are asked 4 times over the course of a year about trips of 75 or more miles during each 3-month reference period. The data provide a complete picture of long trips for the year, but no questions are asked about shorter trips. The large sample size of the ATS permits analysis of flows of people between states and large metropolitan areas.

An integrated design for the ATS and the NPTS could provide useful data for federal, state, and MPO analysis and planning purposes, including consistent estimates of daily and long-distance household travel patterns, in a more cost-effective manner than two separate surveys, neither of which provides a complete picture of household transportation. A possible design (discussed in Chapter 3 in the context of the ATS alone) would be to conduct an annual survey of a relatively small sample of households to provide national estimates, with the sample augmented periodically to provide estimates for states and large MPOs. Each year's combined survey would ask questions both about daily travel patterns and about longer trips. To make the integration of the ATS and NPTS questionnaires feasible and not unduly burdensome to respondents, the sample could be divided into three groups, with one group of households asked only about daily travel, another group asked only about longer trips, and a third group asked about all trips. (This type of design was in fact used for the Nationwide Personal Transportation Survey in 1972 and 1977.)

ADDRESSING GAPS IN HOUSEHOLD TRAVEL DATA

Considering all of the existing household surveys (ATS, NPTS, decennial census, CEX, and RTECS), there are gaps in the data they provide. One such gap is data on commuting. Each of the existing surveys offers data that are relevant to commuting patterns. The decennial census long-form sample makes it possible once every 10 years to map commuting flows among small geographic areas and (since 1980) to determine travel time to work. The NPTS provides updates at 5-year intervals of modes of commuting and distance and time to work, but the sample size permits only limited geographic analysis. The ATS has a larger sample but only covers commuting trips of 75 or more miles (one-way) and does not ask about commuting time. None of these sources provides direct estimates of commuting costs (or about the costs of non-work-related transportation).

The omission of cost information seems quite important, given the productivity implications of commuting time and the expenses incurred by workers. As discussed in Chapter 3, BTS made a deliberate decision to exclude cost data from the ATS on the grounds that households underreport transportation costs. It expects that the U.S. Travel Data Center will develop model-based estimates of long-distance trip costs on the basis of trip characteristics. However, direct survey reports of costs could be useful input to model-based estimates and for validation.

The RTECS asked about modes of commuting and obtained data that permit a rough calculation of the costs of commuting for people who drive. The CEX obtains detailed cost data on transportation for vehicles and trips and usual monthly expenses for public transportation used for work and other purposes. However, the CEX has no data on vehicle miles traveled to work or total vehicle miles, and hence there is no ready way to calculate commuting costs for workers who drive. In addition, there is no way to relate public transportation costs to distance or time traveled. Both the RTECS and CEX sample sizes are quite small, limiting geographic analysis.

In summary, it is not possible to obtain from these data sources a complete picture of commuting flows, times, distances, and costs. The lack of complete data on commuting is an example of a data gap that is likely important to fill for transportation policy planning and analysis. Periodic reviews by BTS of existing transportation data systems, assessed against BTS's vision of user requirements, can identify data gaps and opportunities for data linkages that are important to address in order to serve priority data needs.